A model for isolation with migration avoiding explicit migration events

Graham Jones

2022-10-31, March 8, 2024

art@gjones.name, www.indriid.com

THIS IS PRELIMINARY

1 Introduction

This article describes a model for isolation and migration. In DENIM and IMa3 (Hey et al. (2018)), at every time t, every branch of every gene tree existing at t is assigned to a branch in the species tree, and migration events (times, source and destination branches) are parameters in the model which are sampled during the MCMC sampling. The model here assigns gene tree nodes (coalescences) to species tree branches, and uses an approximation to integrate out the migration events analytically.

My current name for the model is LUCAS = Lineages Unassigned and Coalescences Assigned to Species

1.1 Notation

(This is mainly for organizing names; full definitions in text.)

- G gene tree topology and node (coalescence) times
- Λ parameters assigning coalescences to species tree branches
- never name the species tree?
- m migration rate.
- θ_b population parameter in branch b
- t time measured backwards
- a, b, c, d indices for species tree branches
- s number of current branches
- $\mathcal{L}_b(t)$ lineages in branch b at time t.
- $\mathcal{L}(t)$ all lineages at time t.
- i, j, k, l indices for lineages (=gene tree branches)

- $\breve{A}_i$ event that lineage i goes from start branch and time to end branch and time
- $\hat{t}(i)$, $\breve{t}(i)$ start and end time of lineage i.
- $\hat{b}(i)$, $\breve{b}(i)$ start and end branch of lineage i.
- $\widehat{P}_{bi}(t)$ prob that i is in b at t.
- α average coalescence rate.
- Φ a stochastic rate matrix for states coalesced, together, apart.
- $\mathsf{path}(i)$ path rootwards from lineage i.
- $T_{ij}(t)$ event that $\mathsf{path}(i)$ and $\mathsf{path}(j)$ do not coalesce before t
- $T_{ij}(k,t)$ event that neither $\mathsf{path}(i)$ and $\mathsf{path}(k)$ nor $\mathsf{path}(j)$ and $\mathsf{path}(k)$ coalesce before t
- N_i number of migrations in lineage i.
- $\mathsf{tmrca}(i,j)$ time of the most recent common ancestor of i and j.
- $\mathsf{allS}(t_1,t_2)$ part of species tree that exists during interval $[t_1,t_2]$.
- $\mathsf{ancS}(i)$ branch $\hat{b}(i)$ and branches ancestral to it.
- $\mathsf{descS}(i)$ branch $\breve{b}(i)$ and branches and descending from it.
- $\mathsf{migl}(X)$ total migration rate from a branch segment X of species tree.
- $\mathrm{uexp}(x)$ function $(1-\exp(-x))/x$

1.2 Outline

Time t is measured backwards from present which is at $t=0$. We focus on a single gene tree G. At $t=0$, the gene tree tips are assigned to species tree tips. At each coalescence in G, a parameter (to be estimated) assigns the coalescence to one of the contemporaneous species tree branches. Call the collection of these parameters Λ. Thus, given Λ and the assignments at $t=0$, the probability that a lineage i is in a species tree branch b is known as 0 or 1 at the start and end of every lineage. We denote the start time of i as $\hat{t}(i)$, the end time as $\breve{t}(i)$, the start branch as $\hat{b}(i)$ and the end branch as $\breve{b}(i)$.

The migration during intervals between speciations is modeled by an $s \times s$ rate matrix M, where s is the number of species during the interval, and where M_{bd} is the rate at which a lineage migrates from species tree branch b to species tree branch d. Migration is regarded as going backwards in time, so this is the rate from b at smaller t to d at larger t.

We decompose the gene tree into coalescences where each coalescence 'owns' the two child lineages. For a coalescence between lineages i and j at time $t=\breve{t}(i)=\breve{t}(j)$ in branch $b=\breve{b}(i)=\breve{b}(j)$, we find the probability that i and j do not coalesce before t, and that both are in branch b at time t. Then we deal with other lineages k that exist at time t, and find the probability that they do not coalesce with i or j, given that both are in branch b at time t. For each pair of lineages considered, we split into several cases based on the number of migrations. The cases for very few migrations are handled exactly, and an approximation is used otherwise. Finally we obtain a density for the coalescence time, by multiplying by θ_b^{-1} where θ_b is the usual population size parameter for branch b (long-term effective population times ploidy times mutation rate).

Let $\mathcal{L}(t)$ be the set of all lineages in G existing at t, and $\mathcal{L}_b(t)$ be the set of lineages in branch b at time t. Let $\mathsf{path}(i)$ be the path in G from the start of i to the root. Let $\mathsf{tmrca}(i,j)$ be the time of the most recent common ancestor of i and j, where $\mathsf{path}(i)$ and $\mathsf{path}(j)$ coalesce.

1.3 Comparisons

Comparison with Palczewski and Beerli (2013). There are two sources of inaccuracy in model of Palczewski and Beerli (2013). The first is the lack of independence between probabilities that different lineages are in a branch b at some time t. Consider two species branches b and c with equal populations sizes, and equal migration rates m each way between them. Assume b and c do not merge for a very long time, and suppose that i is assigned to b and j to c at $t = 0$ and the first coalescence is between i and j. Starting from $t = 0$ the lineages behave independently, but once i and j have coalesced to form k, it is only known that $\Pr(k \in \mathcal{L}_b(t)) = \Pr(k \in \mathcal{L}_c(t)) = 1/2$ and the coalescent intensity between k and any other lineage l is $1/(2\theta)$, and expected time to coalescence equal to 2θ. The true situation is that k and l are either together, with initial intensity $1/\theta$ and the model gives an expected time to coalescence larger than θ, or they are apart, with initial intensity 0, and expected time to coalescence larger than $1/(2m)$ since a migration must happen before the coalescence. Overall the expected time to coalescence is larger than $\theta/2 + 1/(4m)$, and for $m \ll 1/\theta$, this may be much larger than 2θ.

The second problem is that even when the lineages behave independently, as they do until the first coalescence, the method overestimates the coalescent intensity. The problem is especially bad when $m \ll 1/\theta$, for example $m = 1, 1/\theta = 10000$. With i, j, b, and c as above, after a time 0.005, the probability that a migration has happened is about 0.01, and the coalescent intensity in the model is about 100 (and growing), producing an expected time to coalescence of less than 0.015. The true value is over 0.5.

The quality of the approximation is discussed in more detail in Palczewski and Beerli (2013). The model is suited to 'populations' with considerable migration between them, but not suitable for 'species' with small rates of migration. The method here resolves the independence problem by introducing the parameters Λ. This gives a 'fresh start' after each coalescence. The approximation is then improved by considering the cases of 0 or 1 migrations within each pair of coalescing lineages separately, and only using an approximation (different from that of Palczewski and Beerli (2013) but in the same spirit) for the case of at least 2 migrations.

Comparison with Hey et al. (2018) (IMa3). This models migrations explicitly. Exact, but presumably slow with lots of migrations. Allows population size parameters to be integrated out analytically, which is not possible in model proposed here. TODO.

Comparison with Flouri et al. (2019). TODO.

2 Decomposition of $\Pr(G)$

2.1 Coalescences

Suppose that i and j are lineages and that for each of them, the time and branch of their start is known. Let $t = \breve{t}(i) = \breve{t}(j)$ and $b = \breve{b}(i) = \breve{b}(j)$. Let $T_{ij}(t)$ be the event that $\mathsf{path}(i)$ and $\mathsf{path}(j)$ do not coalesce before t. If k is a lineage other than i or j, let $T_{ij}(k, t)$ be the event that neither $\mathsf{path}(i)$ and $\mathsf{path}(k)$ nor $\mathsf{path}(j)$ and $\mathsf{path}(k)$ coalesce before t. Denote by $\breve{A}_i$ the event that $\mathsf{path}(i) \in \mathcal{L}_b(t)$, that is, that i reaches the right branch at the right time to coalesce, and similarly for $\breve{A}_j$.

First we find the probability of the events $T_{ij}(t)$, $\breve{A}_i$, and $\breve{A}_j$. Note that this does not depend on whether any other lineages coalesce with $\mathsf{path}(i)$ or $\mathsf{path}(j)$ before t. Secondly we find the probability of the event $T_{ij}(k, t)$, given the coalescence of $\mathsf{path}(i)$ and $\mathsf{path}(j)$. We call i and j the 'coalescers' and other lineages the 'persisters'.

The density for t is then found by multiplying by θ_b^{-1}. We can then 'forget' about i and j and continue with the rest of (G, Λ). Note that lineages in $\mathcal{L}(t) \setminus \{i, j\}$ may start at any time in $[0, t)$ and the probability

that they do not coalesce amongst themselves before t is calculated when dealing with later coalescences.

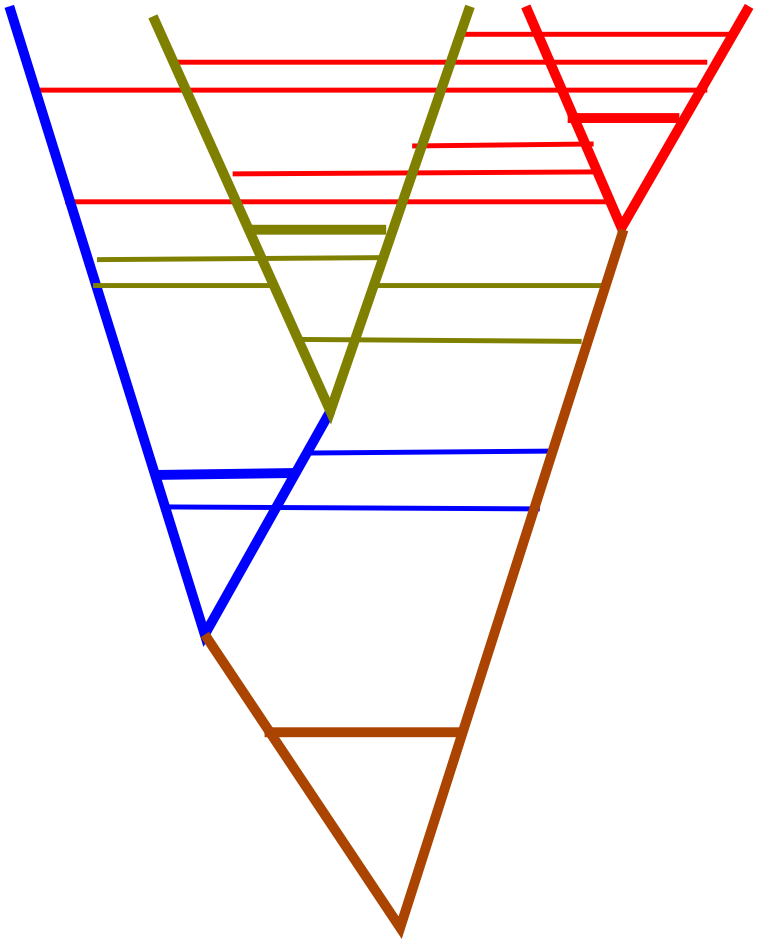

Figure 1: Decomposition of G into 4 coalescences in different colours. At each coalescence, the horizontal lines indicate the pairs of lineages between which potential coalescences are considered.

2.2 Migration counts

We focus on one coalescence and omit the dependence on assignments of previous coalescences to branches. Let N_x be the number of migrations that a lineage x contains before t. (This is all migrations for i and j, but not the ones after t for k.) For the 'coalescers',

$$\begin{aligned}
&\Pr\left(T_{ij}(t) \wedge \breve{A}_i \wedge \breve{A}_j)\right) = \\
&\Pr\left(T_{ij}(t) \wedge \breve{A}_i \wedge \breve{A}_j \wedge (\mathsf{N}_i, \mathsf{N}_j) = (0,0)\right) \; + \\
&\Pr\left(T_{ij}(t) \wedge \breve{A}_i \wedge \breve{A}_j \wedge (\mathsf{N}_i, \mathsf{N}_j) = (0,1)\right) \; + \\
&\Pr\left(T_{ij}(t) \wedge \breve{A}_i \wedge \breve{A}_j \wedge (\mathsf{N}_i, \mathsf{N}_j) = (1,0)\right) \; + \\
&\Pr\left(T_{ij}(t) \wedge \breve{A}_i \wedge \breve{A}_j \wedge \mathsf{N}_i + \mathsf{N}_j \geq 2\right).
\end{aligned} \tag{1}$$

Using the fact that i and j are independent,

$$\begin{aligned}
&\Pr\Big(T_{ij}(t) \wedge \breve{A}_i \wedge \breve{A}_j)\Big) = \\
&\Pr\Big(T_{ij}(t) \mid \breve{A}_i \wedge \breve{A}_j \wedge (\mathsf{N}_i, \mathsf{N}_j) = (0,0)\Big) \Pr(\breve{A}_i \wedge \mathsf{N}_i = 0) \Pr(\breve{A}_j \wedge \mathsf{N}_j = 0) + \\
&\Pr\Big(T_{ij}(t) \mid \breve{A}_i \wedge \breve{A}_j \wedge (\mathsf{N}_i, \mathsf{N}_j) = (0,1)\Big) \Pr(\breve{A}_i \wedge \mathsf{N}_i = 0) \Pr(\breve{A}_j \wedge \mathsf{N}_j = 1) + \\
&\Pr\Big(T_{ij}(t) \mid \breve{A}_i \wedge \breve{A}_j \wedge (\mathsf{N}_i, \mathsf{N}_j) = (1,0)\Big) \Pr(\breve{A}_i \wedge \mathsf{N}_i = 1) \Pr(\breve{A}_j \wedge \mathsf{N}_j = 0) + \\
&\Pr\Big(T_{ij}(t) \wedge \breve{A}_i \wedge \breve{A}_j \wedge \mathsf{N}_i + \mathsf{N}_j \geq 2\Big).
\end{aligned} \tag{2}$$

For the 'persisters', we want

$$\Pr\Big(T_{ij}(k,t) \mid \breve{A}_i \wedge \breve{A}_j\Big) \;=\; \Pr\Big(T_{ij}(k,t) \wedge \breve{A}_i \wedge \breve{A}_j\Big) \;/\; \Pr(\breve{A}_i \wedge \breve{A}_j)$$

and the numerator on the right hand side can be expanded as

$$\begin{aligned}
&\Pr\Big(T_{ij}(k,t)) \wedge \breve{A}_i \wedge \breve{A}_j\Big) \;= \\
&\Pr\Big(T_{ij}(k,t) \wedge \breve{A}_i \wedge \breve{A}_j \wedge (\mathsf{N}_i, \mathsf{N}_j, \mathsf{N}_k) = (0,0,0)\Big) + \\
&\Pr\Big(T_{ij}(k,t) \wedge \breve{A}_i \wedge \breve{A}_j \wedge (\mathsf{N}_i, \mathsf{N}_j, \mathsf{N}_k) = (0,0,1)\Big) + \\
&\Pr\Big(T_{ij}(k,t) \wedge \breve{A}_i \wedge \breve{A}_j \wedge (\mathsf{N}_i, \mathsf{N}_j, \mathsf{N}_k) = (0,1,0)\Big) + \\
&\Pr\Big(T_{ij}(k,t) \wedge \breve{A}_i \wedge \breve{A}_j \wedge (\mathsf{N}_i, \mathsf{N}_j, \mathsf{N}_k) = (1,0,0)\Big) + \\
&\Pr\Big(T_{ij}(k,t) \wedge \breve{A}_i \wedge \breve{A}_j \wedge (\mathsf{N}_i, \mathsf{N}_j, \mathsf{N}_k) = (1,1,0)\Big) + \\
&\Pr\Big(T_{ij}(k,t) \wedge \breve{A}_i \wedge \breve{A}_j \wedge (\mathsf{N}_i + \mathsf{N}_k \geq 2 \vee \mathsf{N}_j + \mathsf{N}_k \geq 2)\Big)
\end{aligned} \tag{3}$$

The first five terms can be expressed like

$$\begin{aligned}
\Pr\Big(T_{ij}(k,t) \wedge \breve{A}_i \wedge \breve{A}_j \wedge (\mathsf{N}_i, \mathsf{N}_j, \mathsf{N}_k) = (n_i, n_j, n_k)\Big) \;= \\
\Pr\Big(T_{ij}(k,t) \;\Big|\; \breve{A}_i \wedge \breve{A}_j \wedge (\mathsf{N}_i, \mathsf{N}_j, \mathsf{N}_k) = (n_i, n_j, n_k)\Big) \;\times \\
\Pr\Big(\breve{A}_i \;\wedge \mathsf{N}_i = n_i\Big) \;\Pr\Big(\breve{A}_j \wedge \mathsf{N}_j = n_j\Big) \;\Pr\left(\mathsf{N}_k = n_k\right)
\end{aligned} \tag{4}$$

In the next section we deal with 'few' migrations, that is, the first three terms in equation (2) and the first five in equation (3). The following section deals with the remaining two terms.

3 Few migrations

3.1 Probabilities of migration counts

Let $\mathsf{allS}(t_1, t_2)$ be the part of the species tree that exists between times t_1 and t_2. For a lineage x, let $\mathsf{ancS}(x)$ be the part of the species tree that is ancestral to $\hat{b}(x)$ at $\hat{t}(x)$. Let $\mathsf{descS}(x)$ be the part of the species tree that is descendant to $\breve{b}(x)$ at $\breve{t}(x)$. Thus for a 'coalescer' $l \in \{i, j\}$, $\mathsf{ancS}(l)$ is the part of the

species tree that $\mathsf{path}(l)$ can reach without migrating, and $\mathsf{descS}(l)$ is the part from which b at t can be reached without migration.

When calculating terms for equation (2) the relevant coalescent times are at $\hat{t}(i)$, $\hat{t}(j)$, and t, and we set $t_0 = \min(\hat{t}(i), \hat{t}(j))$. For equation (3) there is also $\hat{t}(k)$, and we set $t_0 = \min(\hat{t}(i), \hat{t}(j), \hat{t}(k))$. There may also be speciation times within the interval $[t_0, t]$. We divide $\mathsf{allS}(t_0, t)$ into intervals at each of the relevant coalescences, and at every speciation time. Within each interval, the parts of branches that occur will be called 'branch segments' (see Figure 2). The migration rates from a branch segment X to other contemporaneous branch segments are constant, and their sum, multiplied by the duration of X, is the total migration intensity of X, denoted by $\mathsf{migI}(X)$. The total migration intensity for a lineage x (= i, j, or k) is

$$F_{mig}(x) = \sum\{\mathsf{migI}(X) : X \in \mathsf{ancS}(x) \cap \mathsf{allS}(\hat{t}(x), t)\} \tag{5}$$

The distribution of counts follows a Poisson distribution, so we have

$$\Pr(\mathsf{N}_x = 0) = \exp\left[-F_{mig}(x)\right] \tag{6}$$

and

$$\Pr(\mathsf{N}_x = 1) = F_{mig}(x) \exp\left[-F_{mig}(x)\right]. \tag{7}$$

Then $\Pr(\mathsf{N}_i + \mathsf{N}_j \geq 2)$ can be found by subtracting the three other probabilities from 1:

$$\begin{aligned} \Pr(\mathsf{N}_i + \mathsf{N}_j \geq 2) = 1 \ &- \ \Pr(\mathsf{N}_i = 0) \Pr(\mathsf{N}_j = 0) \ - \\ &\Pr(\mathsf{N}_i = 0) \Pr(\mathsf{N}_j = 1) \ - \ \Pr(\mathsf{N}_i = 1) \Pr(\mathsf{N}_j = 0) \\ &= 1 - (1 + F_{mig}(i) + F_{mig}(j)) \exp(-F_{mig}(i) - F_{mig}(j)) \end{aligned} \tag{8}$$

Likewise, $\Pr(\mathsf{N}_i + \mathsf{N}_k \geq 2 \vee \mathsf{N}_j + \mathsf{N}_k \geq 2)$ can be found by subtracting the five other joint probabilities $\Pr((\mathsf{N}_i, \mathsf{N}_j, \mathsf{N}_k) = (n_i, n_j, n_k))$ from 1.

3.2 Probabilities for arrival and migration counts

Let $l \in \{i, j\}$ be a 'coalescer'. If $\mathsf{ancS}(l) \cap \mathsf{allS}(\hat{t}(l), t)$ and $\mathsf{descS}(l) \cap \mathsf{allS}(\hat{t}(l), t)$ are disjoint, then $\Pr(\breve{A}_l | \mathsf{N}_l = 0) = 0$. Otherwise $\mathsf{ancS}(l) \cap \mathsf{allS}(\hat{t}(l), t)$ must be contained in $\mathsf{descS}(l) \cap \mathsf{allS}(\hat{t}(l), t)$, and $\Pr(\breve{A}_l | \mathsf{N}_l = 0) = 1$. then $\Pr(\breve{A}_l \wedge \mathsf{N}_l = 0) = \Pr(\breve{A}_l | \mathsf{N}_l = 0) \Pr(\mathsf{N}_l = 0)$ can be found using equation (6).

The probability that l arrives in b at t given one migration, is the total migration intensity to a segment from where it can reach b at t without further migrations, divided by the total migration intensity to any available segment. Thus

$$\Pr(\breve{A}_l | \mathsf{N}_l = 1) = \frac{\sum\{\mathsf{migI}(X) : X \in (\mathsf{descS}(l) \cap \mathsf{allS}(\hat{t}(l), t)) \setminus \mathsf{ancS}(l)\}}{\sum\{\mathsf{migI}(X) : X \in \mathsf{allS}(\hat{t}(l), t) \setminus \mathsf{ancS}(l)\}}$$

from which $\Pr(\breve{A}_l \wedge \mathsf{N}_l = 1) = \Pr(\breve{A}_l | \mathsf{N}_l = 1) \Pr(\mathsf{N}_l = 1)$ can be found using equation (7).

3.3 The probability of no coalescences

The 'coalescers' case with $\mathsf{N}_i = 0 \wedge \mathsf{N}_j = 0$, and the 'persisters' case with $\mathsf{N}_i = 0 \wedge \mathsf{N}_k = 0$ or $\mathsf{N}_i = 0 \wedge \mathsf{N}_k = 0$ are straightforward, since the location of $\mathsf{path}(i)$, $\mathsf{path}(j)$, and $\mathsf{path}(k)$ is known at all times. During the times that they are together in a branch c, the coalescent intensity is θ_c^{-1}. (It's like a bit of a standard multispecies coalescent calculation.)

For the case of one migration we need a little lemma. Suppose that exactly one migration of a lineage x occurs during some interval of length w during which a branch c exists, that lineage y is in c during the

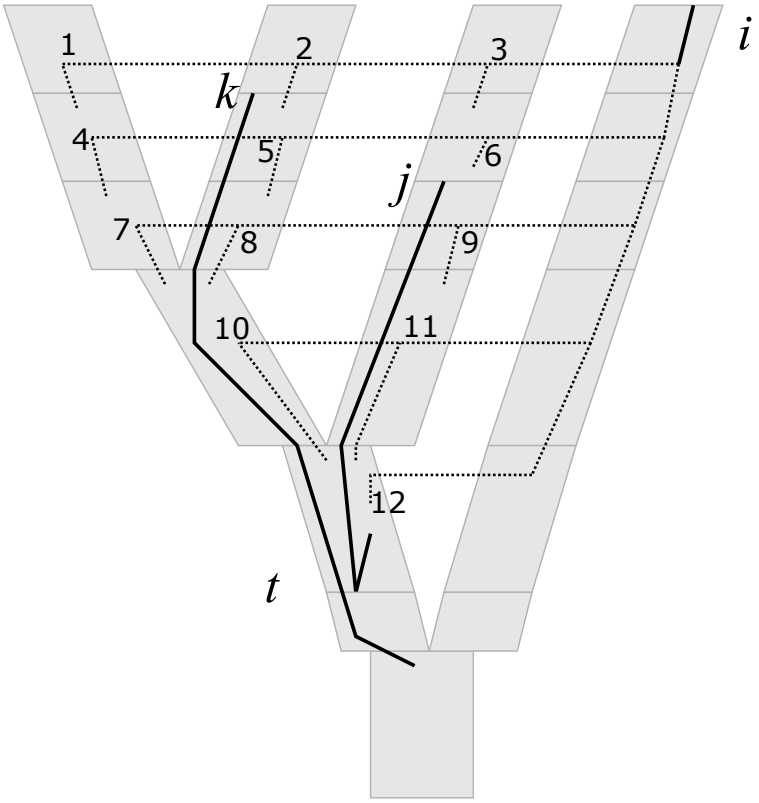

Figure 2: An example for coalescers i and j and persister k where i has one migration, and j and k have none. The branch segments that are possible destinations for i are numbered 1-12.

interval, and that the migration of x is either from or to c. For convenience we define the function $\mathrm{uexp}(x) = (1 - \exp(-x))/x$. Then the probability that there is no coalescence of x and y during the interval is

$$\frac{1 - \exp(-w\theta_c^{-1})}{w\theta_c^{-1}} = \mathrm{uexp}(w\theta_c^{-1}). \tag{9}$$

This can be shown by using the fact that the migration time is uniformly distributed within the interval, and integrating it out. The result is the same whether the time that x and y spend together is at the start or end of the interval.

The method for a single migration is to consider each branch segment in $\mathsf{allS}(t_0, t)$ as a possible destination for the migration. The probability that X' is the destination is found by dividing $\mathsf{migl}(X')$ by the total of the $\mathsf{migl}(X)$'s over the available destinations. We will use '$\to$' to mean 'migrates to', so $x \to X$ means lineage x migrates to X.

For a 'coalescer' l we calculate the probability that the migration went to each branch segment, given $(\breve{A}_i \wedge \breve{A}_j)$. This is

$$\Pr(l \to X' \mid \breve{A}_i \wedge \breve{A}_j) = \Pr(l \to X' \mid \breve{A}_l) = \frac{\mathsf{migl}(X')}{\sum\{\mathsf{migl}(X) : X \in (\mathsf{descS}(l) \cap \mathsf{allS}(\hat{t}(l), t)) \setminus \mathsf{ancS}(l)\}} \tag{10}$$

Figure 2 shows an example. Then, conditioning on each branch segment X', we can find the probability of no coalescence before the segment and the probability of no coalescence after the segment (since we know where both lineages are before and after the segment), and (when needed) use equation (9) for the segment itself.

The 'persisters' case is similar. For the $(\mathsf{N}_i, \mathsf{N}_j, \mathsf{N}_k) = (1, 1, 0)$ case, lineages i and j behave independently, so we can deal with them one at a time and multiply. For the case $N_k = 1$, we do not condition on where k may be at t, and k behaves independently of i and j, so

$$\Pr(k \to X' \mid \breve{A}_i \wedge \breve{A}_j) = \Pr(k \to X') = \frac{\mathsf{migl}(X')}{\sum\{\mathsf{migl}(X) : X \in \mathsf{allS}(\hat{t}(k), t) \setminus \mathsf{ancS}(l)\}} \tag{11}$$

The calculation in this section is described in detail for the case of a fixed number of species in section 5.

4 At least two migrations

Now we turn to the approximation used for more migrations.

4.1 Probability that a lineage is in a given branch at a given time

The probability that a lineage x is in a branch c at a time t is denoted by $\widehat{P}_{cx}(t)$. It is used in section 4.2 to find the probability that two lineages are in the same branch when the second one starts; to find $\Pr(\breve{A}_i)$ and $\Pr(\breve{A}_j)$ (equation 16); and to approximate the probability that i and j are together in branch b given they are together in some branch (equation 17).

Consider a single lineage x in G. Suppose $w \in [\hat{t}(x), \breve{t}(x)]$ and that c is a branch existing at w. Let $\widehat{P}_{cx}(w) = \Pr(\mathsf{path}(x) \in \mathcal{L}_c(w) \mid \mathsf{path}(x) \in \mathcal{L}_{\hat{c}(x)}(\hat{t}(x)))$. Suppose $\hat{t}(x) \le v \le w \le \breve{t}(x)$, and that no speciations occur during $[v, w]$. Then the behaviour of x is determined by the migration matrix M during $[v, w]$ so

$$\Pr(\mathsf{path}(x) \in \mathcal{L}_c(w) \mid \mathsf{path}(x) \in \mathcal{L}_d(v)) = [\exp(M(w - v))]_{dc}. \tag{12}$$

If branches c and d merge to form branch a at time w, we have

$$\widehat{P}_{ax}(w) = \widehat{P}_{cx}(w) + \widehat{P}_{dx}(w). \tag{13}$$

Using these equations we can calculate $\widehat{P}_{cx}(w)$ at any time $w \in [\hat{t}(x), \breve{t}(x)]$, and for any branch c that exists at w. The calculation starts with the assignment $\hat{b}(x)$ at $\hat{t}(x)$. In particular, we can find $\Pr(\breve{A}_i) = \widehat{P}_{bi}(t)$ and $\Pr(\breve{A}_j) = \widehat{P}_{bj}(t)$.

4.2 Approximate Markov model for at least 2 migrations

We use a three state Markov process. The three states are (1) 'coalesced', (2) 'together' (both lineages in same branch, but not coalesced) and (3) 'apart'. The approximation does not account for different population sizes in different branches, nor differing migration rates between contemporaneous branches, but instead uses overall averages. During an interval when there are s branches, the coalescence rate can be approximated by the mean $\alpha = s^{-1}(\sum_b \theta_b^{-1})$. The migration rate m can be apprpoximated by the mean of the off-diagonal entries in M. that is, $m = (s(s-1))^{-1} \sum_{1 \le c \ne d \le s} M_{cd}$.

The process is determined by the following stochastic rate matrix for two paths such as $\mathsf{path}(i)$ and $\mathsf{path}(j)$.

$$\Phi = \begin{bmatrix} 0 & 0 & 0 \\ \alpha & -(\alpha + 2(s-1)m) & 2(s-1)m \\ 0 & 2m & -2m \end{bmatrix}. \tag{14}$$

Given the probabilities that $\mathsf{path}(i)$ and $\mathsf{path}(j)$ are together or apart at the start of an interval of duration u, then the probability that they are in a particular state the end of the period can be found from $\exp(\Phi u)$. For example $\exp(\Phi u)_{32}$ is the probability that $\mathsf{path}(i)$ and $\mathsf{path}(j)$ are together at time u, given they were apart at time 0, and $\exp(\Phi u)_{32} + \exp(\Phi u)_{33}$ is the probability that they have not coalesced by time u. Let $Y_{ij}(u)$ be the state (1,2, or 3) of lineages i and j at time u.

Suppose $\hat{t}(i) \le \hat{t}(j)$, that is, that i starts first. During the interval $[\hat{t}(i), \hat{t}(j)]$, we can find the probability that i arrives in $\hat{b}(j)$ by $\hat{t}(j)$, using $\hat{P}_{bi}(t)$ from section 4.1. Likewise if $\hat{t}(j) \le \hat{t}(i)$.

At a speciation at u, where s branches become $s-1$, the value of $\Pr(Y_{ij}(u) = 2)$ just after the speciation is found from the value of $\Pr(Y_{ij}(u) = 2) + (s(s-1)/2)^{-1} \Pr(Y_{ij}(u) = 3)$ just before the speciation. Likewise the value $\Pr(Y_{ij}(u) = 3)$ just after the speciation is found from the value of $\Pr(Y_{ij}(u) = 2) - (s(s-1)/2)^{-1} \Pr(Y_{ij}(u) = 3)$ just before. Φ can be exponentiated analytically.

4.3 The probability of no coalescences: at least 2 migrations

For the 'coalescers' case, we use Φ to approximate the probability that i and j are together but not coalesced by time t, and multiply by

$$\frac{\widehat{P}_{bi}(t)\widehat{P}_{bj}(t)}{\sum_c \widehat{P}_{ci}(t)\widehat{P}_{cj}(t)} \tag{15}$$

to find the probability that i and j are in b, given that they are together, so

$$\begin{aligned} &\Pr\left(T_{ij}(t) \wedge \breve{A}_i \wedge \breve{A}_j \wedge \mathsf{N}_i + \mathsf{N}_j \ge 2\right) \; = \\ &\Pr\left(T_{ij}(t) \wedge \breve{A}_i \wedge \breve{A}_j \mid \mathsf{N}_i + \mathsf{N}_j \ge 2\right) \; \Pr(\mathsf{N}_i + \mathsf{N}_j \ge 2) \; \simeq \\ &\frac{\widehat{P}_{bi}(t)\widehat{P}_{bj}(t)}{\sum_c \widehat{P}_{ci}(t)\widehat{P}_{cj}(t)} \Pr(Y_{ij}(t) = 2) \; \Pr(\mathsf{N}_i + \mathsf{N}_j \ge 2) \end{aligned} \tag{16}$$

For the 'persisters' case, we assume approximate independence of $(\breve{A}_i \wedge \breve{A}_j)$ and the other events and estimate

$$\begin{aligned}
&\Pr\left(T_{ij}(k,t) \wedge \breve{A}_i \wedge \breve{A}_j \wedge (\mathsf{N}_i + \mathsf{N}_k \geq 2 \vee \mathsf{N}_j + \mathsf{N}_k \geq 2)\right) \simeq \\
&\Pr\left(T_{ij}(k,t) \wedge (\mathsf{N}_i + \mathsf{N}_k \geq 2 \vee \mathsf{N}_j + \mathsf{N}_k \geq 2)\right) \Pr(\breve{A}_i \wedge \breve{A}_j) = \\
&\Pr\left(T_{ij}(k,t) \mid (\mathsf{N}_i + \mathsf{N}_k \geq 2 \vee \mathsf{N}_j + \mathsf{N}_k \geq 2)\right) \Pr(\mathsf{N}_i + \mathsf{N}_k \geq 2 \vee \mathsf{N}_j + \mathsf{N}_k \geq 2) \Pr(\breve{A}_i \wedge \breve{A}_j) = \\
&\Pr\left(T_{ij}(k,t) \mid (\mathsf{N}_i + \mathsf{N}_k \geq 2 \vee \mathsf{N}_j + \mathsf{N}_k \geq 2)\right) \Pr(\mathsf{N}_i + \mathsf{N}_k \geq 2 \vee \mathsf{N}_j + \mathsf{N}_k \geq 2) \widehat{P}_{bi}(t) \widehat{P}_{bj}(t)
\end{aligned} \tag{17}$$

and then assume approximate independence of $(\mathsf{tmrca}(i,k) \geq t)$ and $(\mathsf{tmrca}(j,k) \geq t)$ given the condition on counts, and use Φ to estimate these, so that

$$\begin{aligned}
&\Pr\left(T_{ij}(k,t) \mid (\mathsf{N}_i + \mathsf{N}_k \geq 2 \vee \mathsf{N}_j + \mathsf{N}_k \geq 2)\right) \simeq \\
&(\Pr(Y_{ik}(t) = 2) + \Pr(Y_{ik}(t) = 3)) \, (\Pr(Y_{jk}(t) = 2) + \Pr(Y_{jk}(t) = 3))
\end{aligned} \tag{18}$$

5 Details for fixed number of species and simple migration matrix

In this section, we restrict to the case where the migration rate is the same between any pair of contemporaneous species tree branches. It is quite close to being pseudo-code, and is implemented in the tests in section 6.

we have $M_{bd} = m$ for every $b \neq d$, and $M_{bb} = -(s-1)m$. Let I be the $s \times s$ identity matrix and U be an $s \times s$ matrix filled with $1/s$. Then $M = smU - smI$ and it is straightforward to show that for any real number x, we have $\exp(Mx) = U + (I - U)e^{-smx}$.

Suppose there are s species at all times - a simple island model, no tree. I use $\mathbb{I}[]$ as the indicator function, equal to one if its argument is true, else zero. We find values for equation (2) then equation (3). Let $b = \breve{b}(i) = \breve{b}(j)$, and $t = \breve{t}(i) = \breve{t}(j)$.

5.1 'coalescers'

Let $u_i = \max(\hat{t}(j) - \hat{t}(i), 0)$. Let $u_j = \max(\hat{t}(i) - \hat{t}(j), 0)$. Let $w = t - \max(\hat{t}(i), \hat{t}(j))$. Then u_i is the time i spends alone, u_j is the time that j spends alone, and w is the time they spend together.

5.1.1 Migration counts

$$\begin{aligned}
\Pr(\mathsf{N}_i = 0) &= \exp(-(s-1)m(u_i + w)) \\
\Pr(\mathsf{N}_i = 1) &= (s-1)m(u_i + w)\exp(-(s-1)m(u_i + w))
\end{aligned}$$

Likewise

$$\begin{aligned}
\Pr(\mathsf{N}_j = 0) &= \exp(-(s-1)m(u_j + w)) \\
\Pr(\mathsf{N}_j = 1) &= (s-1)m(u_j + w)\exp(-(s-1)m(u_j + w))
\end{aligned}$$

Then

$$\Pr(\mathsf{N}_i + \mathsf{N}_j \geq 2) = 1 - \Pr(\mathsf{N}_i = 0)\Pr(\mathsf{N}_j = 0) - \Pr(\mathsf{N}_i = 0)\Pr(\mathsf{N}_j = 1) - \Pr(\mathsf{N}_i = 1)\Pr(\mathsf{N}_j = 0)$$

5.1.2 Joint probabilities of arrival and counts

$$\Pr(\breve{A}_i \wedge \mathsf{N}_i = 0) \;=\; \Pr(\breve{A}_i | \mathsf{N}_i = 0) \Pr(\mathsf{N}_i = 0) \;=\; \mathbb{I}[\hat{b}(i) = b] \Pr(\mathsf{N}_i = 0)$$
$$\Pr(\breve{A}_i \wedge \mathsf{N}_i = 1) \;=\; \Pr(\breve{A}_i | \mathsf{N}_i = 1) \Pr(\mathsf{N}_i = 1) \;=\; \mathbb{I}[\hat{b}(i) \neq b] \times (s-1)^{-1} \Pr(\mathsf{N}_i = 1)$$

since i must start elsewhere than branch b, and it has $s-1$ branches to go to, one of which is b. Likewise

$$\Pr(\breve{A}_j \wedge \mathsf{N}_j = 0) \;=\; \mathbb{I}[\hat{b}(j) = b] \Pr(\mathsf{N}_j = 0)$$
$$\Pr(\breve{A}_j \wedge \mathsf{N}_j = 1) \;=\; \mathbb{I}[\hat{b}(j) \neq b] \times (s-1)^{-1} \Pr(\mathsf{N}_j = 1).$$

5.1.3 The $(\mathsf{N}_i, \mathsf{N}_j) = (0,0)$ case

When a coalescence is possible in the (0,0) case (that is, when $\hat{b}(i) = \hat{b}(j) = b$),

$$\Pr\Big(T_{ij}(t) \mid \breve{A}_i \wedge \breve{A}_j \wedge (\mathsf{N}_i, \mathsf{N}_j) = (0,0)\Big) = \exp(-w\theta_b^{-1})$$

5.1.4 The $(\mathsf{N}_i, \mathsf{N}_j) = (0,1)$ case

When a coalescence is possible in the (0,1) case (that is, when $\hat{b}(i) = b$ and j migrates to b), i may migrate to b before $\breve{t}(j)$ or after $\breve{t}(j)$, and

$$\Pr\Big(T_{ij}(t) \mid \breve{A}_i \wedge \breve{A}_j \wedge (\mathsf{N}_i, \mathsf{N}_j) = (0,1)\Big) \;=\; (u_j + w)^{-1}\big(u_j \exp(-w\theta_b^{-1}) \;+\; w\mathrm{uexp}(w\theta_b^{-1})\big)$$

5.1.5 The $(\mathsf{N}_i, \mathsf{N}_j) = (1,0)$ case

Same as previous case with i and j swapped.

5.1.6 The $\mathsf{N}_i + \mathsf{N}_j \geq 2$ case

Let $v = |\hat{t}(i) - \hat{t}(j)|$ be the time that only one of i and j exist. Then the probability that i and j are together at $t_{max} = \max(\hat{t}(i), \hat{t}(j))$ is approximately

$$\Pr(Y_{ij}(t_{max}) = 2) = (1/s)(1 - \exp(-smv)) + \exp(-smv)\mathbb{I}[\hat{b}(i) = \hat{b}(j)] \tag{19}$$

Then

$$\Pr(Y_{ij}(t) = 2) = \exp(\Phi w)_{32}(1 - \Pr(Y_{ij}(t_{max}) = 2)) + \exp(\Phi w)_{22} \Pr(Y_{ij}(t_{max}) = 2) \tag{20}$$

and

$$\begin{aligned} &\Pr\Big(\mathsf{tmrca}(i,j) \geq t \wedge \breve{A}_i \wedge \breve{A}_j \wedge (\mathsf{N}_i + \mathsf{N}_j \geq 2)\Big) \;\simeq \\ &\frac{\widehat{P}_{bi}(t)\widehat{P}_{bj}(t)}{\sum_c \widehat{P}_{ci}(t)\widehat{P}_{cj}(t)} \;\Pr(Y_{ij}(t) = 2) \;\Pr(\mathsf{N}_i + \mathsf{N}_j \geq 2) \end{aligned} \tag{21}$$

5.2 'persisters'

Let u_i be the time that i spends alone, before j or k begin, or zero if j or k start first, so $u_i = \max(0, \min(\hat{t}(j), \hat{t}(k)) - \hat{t}(i))$. Likewise, define u_j and u_k. Let v_{ij} be the time during which i and j exist, but k does not, so $v_{ij} = \max(0, \hat{t}(k) - \max(\hat{t}(i), \hat{t}(j)))$. Likewise define v_{ik} and v_{jk}. Only one of u_i, u_j, u_k and one of v_{ij}, v_{ik}, v_{jk} can be nonzero. Let w be the time during which all three exist, so $w = t - \max(\hat{t}(i), \hat{t}(j), \hat{t}(k))$. we also set $t_i = t - \hat{t}(i)$, $t_j = t - \hat{t}(j)$, $t_k = t - \hat{t}(k)$. These are the duration of i and j, but only part of the duration of k.

5.2.1 Migration counts

$$\begin{aligned}\Pr(\mathsf{N}_k = 0) &= \exp(-(s-1)mt_k)\\ \Pr(\mathsf{N}_k = 1) &= (s-1)mt_k \exp(-(s-1)mt_k)\end{aligned}$$

with similar expressions for N_i and N_j, and

$$\begin{aligned}\Pr(\mathsf{N}_i + \mathsf{N}_k \geq 2 \vee \mathsf{N}_j + \mathsf{N}_k \geq 2) &= 1 - \Pr(N_i = 0)\Pr(N_j = 0)\Pr(N_k = 0) -\\ &\Pr(N_i = 0)\Pr(N_j = 0)\Pr(N_k = 1) - \Pr(N_i = 0)\Pr(N_j = 1)\Pr(N_k = 0) -\\ &\Pr(N_i = 1)\Pr(N_j = 0)\Pr(N_k = 0) - \Pr(N_i = 1)\Pr(N_j = 1)\Pr(N_k = 0)\end{aligned}$$

5.2.2 Joint probabilities of arrival and counts

We have $\Pr(\breve{A}_i \wedge \mathsf{N}_i = 0)$, $\Pr(\breve{A}_j \wedge \mathsf{N}_j = 0)$, $\Pr(\breve{A}_i \wedge \mathsf{N}_i = 1)$, and $\Pr(\breve{A}_j \wedge \mathsf{N}_j = 1)$ from section 5.1.2. We do not need anything for k.

5.2.3 The $(\mathsf{N}_i, \mathsf{N}_j, \mathsf{N}_k) = (0,0,0)$ case

This case is impossible unless $\hat{b}(i) = \hat{b}(j) = b$. Given this, if $\hat{b}(k) \neq b$ no coalescence between k and i or j is possible, and if $\hat{b}(k) = b$, it may coalesce during the intervals that k and i or j co-exist. Thus

$$\Pr\left(T_{ij}(k,t) \;\middle|\; \breve{A}_i \wedge \breve{A}_j \wedge (\mathsf{N}_i, \mathsf{N}_j, \mathsf{N}_k) = (0,0,0)\right) = \mathbb{I}[\hat{b}(k) \neq b] + \mathbb{I}[\hat{b}(k) = b]\exp(-(v_{ik} + v_{jk} + 2w)\theta_b^{-1})$$

5.2.4 The $(\mathsf{N}_i, \mathsf{N}_j, \mathsf{N}_k) = (0,0,1)$ case

Again, this case is impossible unless $\hat{b}(i) = \hat{b}(j) = b$. There is no coalescence unless $\hat{b}(k) \neq b$ and k migrates to b. Given $\hat{b}(k) \neq b$, the probability that k migrates to b is $1/(s-1)$. It may arrive in b before i or j have started, when one exists, buit not the other, or after both have started. The probabilities of these three arrival types are equal to the fraction of t_k during which they can occur. Thus

$$\begin{aligned}\Pr\left(T_{ij}(k,t) \;\middle|\; \breve{A}_i \wedge \breve{A}_j \wedge (\mathsf{N}_i, \mathsf{N}_j, \mathsf{N}_k) = (0,0,1)\right) &= \mathbb{I}[\hat{b}(k) = b] +\\ \mathbb{I}[\hat{b}(k) \neq b]\,(s-1)^{-1}\,t_k^{-1} &\times \Big(u_k \exp(-(v_{ik} + v_{jk} + 2w)\theta_b^{-1}) +\\ (v_{ik} + v_{jk})\mathrm{uexp}(u_i + u_j)\theta_b^{-1})\exp(-2w\theta_b^{-1}) &+ w\,\mathrm{uexp}(2w\theta_b^{-1})\Big)\end{aligned}$$

5.2.5 The $(\mathsf{N}_i, \mathsf{N}_j, \mathsf{N}_k) = (0, 1, 0)$ case

This case is impossible unless $\hat{b}(i) = b$ and $\hat{b}(j) \neq b$. If $\hat{b}(k)$ is neither b nor $\hat{b}(j)$, there can be no coalescence.

If $\hat{b}(k) = b$ and $\hat{b}(k) \neq \hat{b}(i)$, then k and i are together for a time $v_{ik} + w$. The lineage j may migrate to b before k starts during an interval of length $u_j + v_{ij}$, or while k exists, during an interval of length $v_{jk} + w$.

If $\hat{b}(k) \neq b$ and $\hat{b}(k) = \hat{b}(j)$, then k cannot coalesce with i, but may coalesce with j before j leaves $\hat{b}(k)$.

Note $\hat{b}(k) = b$ and $\hat{b}(k) = \hat{b}(j)$ cannot happen, since $\hat{b}(j) \neq b$.

$$
\begin{aligned}
&\Pr\Big(T_{ij}(k,t) \;\Big|\; \breve{A}_i \wedge \breve{A}_j \wedge (\mathsf{N}_i, \mathsf{N}_j, \mathsf{N}_k) = (0,1,0)\Big) \;=\; \mathbb{I}[\hat{b}(k) \neq b]\; \mathbb{I}[\hat{b}(k) \neq \hat{b}(j)] \;+ \\
&\mathbb{I}[\hat{b}(k) = b]\; \mathbb{I}[\hat{b}(k) \neq \hat{b}(i)]\; \exp(-(v_{ik} + w)\theta_b^{-1}) \times \\
&\qquad t_j^{-1}\Big((u_j + v_{ij})\exp(-(v_{jk} + w)\theta_b^{-1}) + (v_{jk} + w)\mathrm{uexp}\big((v_{jk} + w)\theta_b^{-1}\big)\Big) \;+ \\
&\mathbb{I}[\hat{b}(k) \neq b]\; \mathbb{I}[\hat{b}(k) = \hat{b}(j)]\; t_j^{-1}\Big((u_j + v_{ij}) + (v_{jk} + w)\mathrm{uexp}\big((v_{jk} + w)\theta_{\hat{b}(k)}^{-1}\big)\Big)
\end{aligned}
$$

5.2.6 The $(\mathsf{N}_i, \mathsf{N}_j, \mathsf{N}_k) = (1, 0, 0)$ case

The same as last subsection 5.2.5 with i and j swapped.

5.2.7 The $(\mathsf{N}_i, \mathsf{N}_j, \mathsf{N}_k) = (1, 1, 0)$ case

This is impossible unless $\hat{b}(i) \neq b$ and $\hat{b}(j) \neq b$. Lineages i and j behave independently, so we can deal with them one at a time and multiply. If $\hat{b}(k) = b$, then i may migrate to b before k starts during an interval of length $u_i + v_{ij}$, or while k exists, during an interval of length $v_{ik} + w$. Similarly for j.

If $\hat{b}(k) \neq b$, then i may migrate away from $\hat{b}(k)$ before k starts during an interval of length $u_i + v_{ij}$, or while k exists, during an interval of length $v_{ik} + w$. Similarly for j.

$$
\begin{aligned}
&\Pr\Big(T_{ij}(k,t) \;\Big|\; \breve{A}_i \wedge \breve{A}_j \wedge (\mathsf{N}_i, \mathsf{N}_j, \mathsf{N}_k) = (1,1,0)\Big) \;= \\
&\Big[\mathbb{I}[\hat{b}(k) = b]\; t_i^{-1}\; \Big((u_i + v_{ij})\exp(-(v_{ik} + w)\theta_b^{-1}) + (v_{ik} + w)\mathrm{uexp}\big((v_{ik} + w)\theta_b^{-1}\big)\Big) \;+ \\
&\mathbb{I}[\hat{b}(k) = \hat{b}(i)] t_i^{-1}\Big(u_i + v_{ij} + (v_{ik} + w)\mathrm{uexp}\big((v_{ik} + w)\theta_{\hat{b}(k)}^{-1}\big)\Big)\Big) \;+ \\
&\mathbb{I}[\hat{b}(k) \neq b]\mathbb{I}[\hat{b}(k) \neq \hat{b}(i)]\Big] \quad \times \Big[\text{same thing with } i \text{ and } j \text{ swapped}\Big].
\end{aligned}
$$

5.2.8 The $\mathsf{N}_i + \mathsf{N}_k \geq 2 \vee \mathsf{N}_j + \mathsf{N}_k \geq 2$ case

We assume approximate independence to obtain

$$
\begin{aligned}
&\Pr\Big(T_{ij}(k,t) \wedge \breve{A}_i \wedge \breve{A}_j \wedge (\mathsf{N}_i + \mathsf{N}_k \geq 2 \vee \mathsf{N}_j + \mathsf{N}_k \geq 2)\Big) \;\simeq \\
&\Pr(\mathsf{tmrca}(i,k) \geq t)\;\; \Pr(\mathsf{tmrca}(j,k) \geq t)\;\; \Pr(\breve{A}_i \wedge \breve{A}_j)\;\; \Pr(\mathsf{N}_i + \mathsf{N}_k \geq 2 \vee \mathsf{N}_j + \mathsf{N}_k \geq 2)
\end{aligned}
$$

and $\Pr(\mathsf{tmrca}(i,k) \geq t)$ and $\Pr(\mathsf{tmrca}(j,k) \geq t)$ can be approximated using Φ.

Let $v = |\hat{t}(i) - \hat{t}(k)| = (u_i + v_{ij} + u_k + v_{jk})$ be the time that only one of i and k exist. The probability that i and k are together at $t_{max} = \max(\hat{t}(i), \hat{t}(k))$ is

$$\Pr(Y_{ik}(t_{max}) = 2) = (1/s)(1 - \exp(-smv) + \exp(-smv)\mathbb{I}[\hat{b}(i) = \hat{b}(k)] \tag{22}$$

Then the probability that i and k have not coalesced by t, namely $\Pr(\mathsf{tmrca}(i,k) \geq t)$, is approximated as

$$\begin{aligned} &\Pr(Y_{ik}(t) = 2) \ + \ \Pr(Y_{ik}(t) = 3) \ = \\ &\exp(\Phi w)_{32}(1 - \Pr(Y_{ik}(t_{max}) = 2)) + \exp(\Phi w)_{22} \Pr(Y_{ik}(t_{max}) = 2) \ + \\ &\exp(\Phi w)_{33}(1 - \Pr(Y_{ik}(t_{max}) = 2)) + \exp(\Phi w)_{23} \Pr(Y_{ik}(t_{max}) = 2) \end{aligned} \tag{23}$$

There is a similar expression for $\Pr(\mathsf{tmrca}(j,k) \geq t)$.

6 Tests of the approximation

Some tests were carried out. They were all for a fixed number of species (like between speciations in a species tree). The coalescence rate was the same in every species and every test at 1000. Thus the expected time for two lineages in the same species to coalesce without migration is 0.001. The migration rate is constant and the same between any pair of species.

For each configuration, gene trees are sampled directly by simulating coalescence and migration events. This is taken to be a sample from the true distribution. There were 1000 trees in each sample.

A MCMC sample of gene trees is generated using the approximate probability described in previous sections. Three MCMC operators were used on the gene tree. One was NNI. Another changes the height of a coalescence. The third chooses a random species for a coalescence. The MCMC run was 1000000, sampled every 1000, to produce 1000 trees. The run started with a single simulated tree, with no burn-in.

The distributions of coalescence times of the direct and MCMC samples are then compared using plots of CDFs for successive intervals between coalescence times. In result files, black is for direct and red for approximation. There are also text files with median and mean coalescence times (not intervals).

The tests come in groups, with the migration rate varying within each group (with values 0.5, 5, 50, 500).There are three groups with 2 lineages. The output results are combined into one file `tests2lin.pdf` for these.

- `Testspp2lins2m05,...,Testspp2lins2m500`:
  2 species, 2 lineages in different species
- `Testspp3lins2m05,...,Testspp3lins2m500`:
  3 species, 2 lineages in different species
- `Testspp6lins2m05,...,Testspp6lins2m50`:
  6 species, 2 lineages in different species

And then three groups with more lineages, one result file each. (`tests3spp5lin320.pdf`, `tests5spp5lin11111.pdf`, and `tests5spp5lin44444.pdf`.)

- `Testspp3lins3+2+0m05,...,Testspp3lins3+2+0m05`:
  3 species, 5 lineages, 3 in one species, 2 in another, 0 in the third
- `Testspp5lins1+1+1+1+1m05,...,Testspp5lins1+1+1+1+1m500`:
  5 species, 5 lineages, one lineage in each species
- `Testspp5lins4+4+4+4+4m05,...,Testspp5lins4+4+4+4+4m500`:
  5 species, 20 lineages, 4 lineages in each species

6.1 Config files

The configuration for each test is specified in a file like this:

```
seeds_GTree_Direct_MCMC 42 43 44
nSpeciesTips 3
mRate 50
cRates 1000.0 1000.0 1000.0
nGtreeTips 3 2 0
```

`mRate` is migration rate, `cRates` are coalescence rates.

References

Tomáš Flouri, Xiyun Jiao, Bruce Rannala, and Ziheng Yang. A Bayesian Implementation of the Multispecies Coalescent Model with Introgression for Phylogenomic Analysis. *Molecular Biology and Evolution*, 37(4):1211–1223, 12 2019. ISSN 0737-4038. doi: 10.1093/molbev/msz296. URL `https://doi.org/10.1093/molbev/msz296`.

Jody Hey, Yujin Chung, Arun Sethuraman, Joseph Lachance, Sarah Tishkoff, Vitor C Sousa, and Yong Wang. Phylogeny Estimation by Integration over Isolation with Migration Models. *Molecular Biology and Evolution*, 35(11):2805–2818, 08 2018. ISSN 0737-4038. doi: 10.1093/molbev/msy162. URL `https://doi.org/10.1093/molbev/msy162`.

Michal Palczewski and Peter Beerli. A continuous method for gene flow. *Genetics*, 194(3):687–696, 2013. ISSN 0016-6731. doi: 10.1534/genetics.113.150904. URL `http://www.genetics.org/content/194/3/687`.